SOME GENERALIZATIONS OF CARATHÉODORY'S THEOREM VIA BARYCENTRES, WITH APPLICATION TO MATHEMATICAL PROGRAMMING

BY S. H. TIJS AND J. M. BORWEIN

ABSTRACT. A theorem on the barycentre of a measure is proven which leads to generalization of Carathéodory's theorem and to extension of various results. A mathematical programming problem is examined in application.

- 1. **Introduction.** In theorem 1 of this paper we present a result concerning the placement of the barycentre of a probability measure, which is used in theorem 2 to generalize Carathéodory's theorem. Theorem 3 then applies this result to CS-closed sets. Theorem 3 is then itself applied to produce an extension of a recent result of Cook [2], which can also be seen as a generalization of Carathéodory's theorem. The final section of this paper considers a very general linear programming application of theorem 2.
- 2. **Notation and preliminaries.** All discussion takes place in the setting of a normed linear space X over \mathbb{R} (although many results can be extended to a more general setting). The space of continuous linear functionals of X is denoted by X^* . If V is an arbitrary subset of X, then V^c is the complement of V in X, conv V is the convex hull of V, cl V is the closure of V, int V is the interior of V and, if C is convex, ri C is the relative interior of C with respect to the smallest closed affine subspace containing C. A measure μ on X will be a non-negative regular Borel measure on the σ -ring generated by the open subsets of X. A probability measure on X is a measure μ with $\mu(X) = 1$. The support supp (μ) of a measure μ on X is defined by

supp
$$(\mu) = \{x \in X : \mu(U) > 0 \text{ for each open neighbourhood } U \text{ of } x\}.$$

The probability measure with mass 1 at x will be written ε_x . The barycentre $b(\mu)$ of a probability measure μ is defined by

$$x^*(b(\mu)) = \int_X x^*(x) d\mu(x)$$
 for all $x^* \subset X^*$

if such a point $b(\mu)$ exists. Further S^{∞} will denote the set of infinite sequences $\{q_n\}$ of non-negative terms with $\sum_{n=1}^{\infty} q_n = 1$. Finally $\mathbb{R} = \mathbb{R} \cup \{-\infty, \infty\}$ and $\pi_i : \mathbb{R}^m \to \mathbb{R}$ is the *i*-th coordinate functional (i = 1, 2, ..., m).

The following proposition will be used.

Received by the editors December 2, 1977 and in revised form March 13, 1978 and December 20, 1978.

Proposition 1. Let μ be a probability measure on X. Then $\mu(\text{supp}(\mu)) = 1$.

Proof. Since $\operatorname{supp}(\mu)$ is closed, this set is measurable and also its complement. Let us suppose that $\operatorname{supp}(\mu)^c$ has positive measure. Then, by the regularity of μ , some compact subset K of $\operatorname{supp}(\mu)^c$ also has positive measure. Now each point in K belongs to an open set of measure zero. Since K is compact some finite number of such open sets covers K. But then K has measure zero, which is impossible. Hence $\operatorname{supp}(\mu)^c$ has measure 0, $\mu(\operatorname{supp}(\mu)) = 1$. \square

3. Central results. Our first theorem says something about the placement of the barycentre of a probability measure.

Theorem 1. Let μ be a probability measure on X with barycentre $b(\mu)$. Suppose V is a subset of X satisfying

$$(3.1) supp(\mu) = cl V,$$

(3.2)
$$ri \operatorname{conv} V \neq \emptyset$$
.

Then $b(\mu) \in ri \text{ conv } V$.

Proof. Suppose that $b(\mu) \not\in ri$ conv V. Let M be the closed affine span of V and suppose, by translation, that M is a subspace. If $b(\mu)$ lies in M, one may separate $b(\mu)$ from the non-empty open set ri conv V in M. The extension to X of the separating functional on M produces some $x^* \in X^*$ with

(3.3)
$$x^*(b(\mu)) < x^*(y)$$
 for each $y \in ri \text{ conv } V$.

In case $b(\mu)$ does not lie in the closed subspace M, (3.3) can still be supposed to hold. It follows from (3.3) that there is a $\bar{v} \in V$ such that $x^*(b(\mu)) < x^*(\bar{v})$. But then there is an open set U containing \bar{v} and some $\varepsilon > 0$ such that

(3.4)
$$x^*(b(\mu)) \le x^*(u) - \varepsilon$$
 for each $u \in U$.

Moreover, $\mu(U) > 0$ because $\bar{v} \in \text{supp}(\mu)$ by (3.1). Also, by (3.3) and (3.1), we have

(3.5)
$$x^*(b(\mu)) \le x^*(s) \quad \text{for each} \quad s \in \text{supp}(\mu).$$

But then, in view of (3.4) and (3.5), we have

$$x^*(b(\mu)) = \int_x x^*(x)d\mu(x) \ge \int_{U \cap \text{supp}(\mu)} x^*(x)d\mu(x)$$

$$+ \int_{U^c \cap \text{supp}(\mu)} x^*(x)d\mu(x) \ge \mu(U \cap \text{supp}(\mu))(x^*(b(\mu)) + \varepsilon)$$

$$+ \mu(U^c \cap \text{supp}(\mu))x^*(b(\mu))$$

$$= \mu(\text{supp}(\mu))x^*(b(\mu)) + \varepsilon\mu(U \cap \text{supp}(\mu)).$$

Since $\mu(\text{supp}(\mu)) = 1$ by proposition 1 and thus $\mu(U \cap \text{supp}(\mu)) = \mu(U) > 0$, we have derived the contradiction that $x^*(b(\mu)) \ge x^*(b(\mu)) + \varepsilon \mu(U)$. Hence $b(\mu) \in ri$ conv V. \square

Now we recall Carathéodory's theorem: Let V be a subset of \mathbb{R}^m and let $c \in \text{conv } V$. Then there exists a finite subset W of V with at most m+1 elements such that $c \in \text{conv } W$.

We note that

(3.6) $c \in \text{conv } V \text{ iff}$ there is a probability measure μ such that

 $\operatorname{supp}(\mu)$ is a finite subset of V and $b(\mu) = c$.

[Suppose $q_i \ge 0$, $v(i) \in V$ (i = 1, ..., s) and $\sum_{i=1}^{s} q_i = 1$. Then the element $c = \sum_{i=1}^{s} q_i v(i)$ in conv V corresponds with the probability measure $\mu = \sum_{i=1}^{s} q_i \, \varepsilon_{v(i)}$ and $b(\mu) = c$.]

Now we can reformulate Carathéodory's theorem as follows: For each probability measure μ on \mathbb{R}^m with a finite support in $V \subseteq \mathbb{R}^m$, there is a probability measure ν on \mathbb{R}^m such that $\operatorname{supp}(\nu)$ is a subset of V with at most m+1 elements and such that the barycentres of μ and ν coincide. In view of this reformulation the following theorem can be seen as a generalization of Carathéodory's theorem.

Theorem 2. Let μ be a probability measure on \mathbb{R}^m such that $\int \pi_i(x) d\mu(x) \in \mathbb{R}$ for i = 1, 2, ..., m. Let V be a subset of \mathbb{R}^m such that $\operatorname{supp}(\mu) = \operatorname{cl}(V)$. Then there exists a probability measure ν on \mathbb{R}^m with

$$b(\mu) = b(\nu) = \left(\int \pi_2(x) \ d\mu(x), \ldots, \int \pi_m(x) \ d\mu(x) \right)$$

and supp(v) is a subset of V with at most m+1 elements.

Proof. It is obvious that $b(\mu)$ exists and that $(b(\mu))_i = \int \pi_i(x) d\mu(x)$ for i = 1, ..., m. Since it is well-known that each non-empty convex set in a finite-dimensional space has a non-empty relative interior and since $\operatorname{supp}(\mu) = \operatorname{cl}(V)$, we can apply theorem 1 and conclude that $b(\mu) \in \operatorname{ri} \operatorname{conv} V \subset \operatorname{conv} V$. In view of (3.6) and the reformulation of Carathéodory's theorem there exists a probability measure ν with the desired properties. \square

Recall that a convex set C is said to be CS-closed (cf. Jameson [7], p. 114) if, for any sequence $\{c(n)\}$ in C and $\{q_n\} \in S^{\infty}$ the convex sum $\sum_{n=1}^{\infty} q_n \ c(n)$ is in C whenever the sum converges. Clearly, closed convex sets are CS-closed. We shall use theorem 1 to provide a proof of the known result that all finite dimensional convex sets are CS-closed.

THEOREM 3. Suppose C is a convex subset of \mathbb{R}^m . Let $\sum_{n=1}^{\infty} q_n \ c(n)$ be a convergent convex sum with $c(n) \in C$ for each $n \in N$. Then $\sum_{n=1}^{\infty} q_n \ c(n)$ is also a convex combination of at most m+1 elements of C. Particularly, C is CS-closed.

Proof. Put $c = \sum_{n=1}^{\infty} q_n \ c(n)$, $V = \{c(n) : q_n > 0\}$ and $\mu = \sum_{n=1}^{\infty} q_n \ \varepsilon_{c(n)}$. Then μ is a probability measure on \mathbb{R}^m with supp $(\mu) = \operatorname{cl} V$, $b(\mu) = c$ and n conv $V \neq \emptyset$. So we may conclude from theorem 1 that $c \in \operatorname{conv} V \subset C$. Hence C is CS-closed. The other conclusion of the theorem follows from Carathéodory's theorem. \square

It is apparent on inspection of the proof of theorem 1 that it continues to hold if (3.1) is replaced by

(3.1)'
$$\operatorname{supp}(\mu) \subset \operatorname{cl} \operatorname{conv} V$$
; $ri \operatorname{conv} V \cap \operatorname{supp}(\mu) \neq \emptyset$,

which is not strictly comparable to (3.1). Consider $V = \{0, 1\} \subset \mathbb{R}$, $\mu = \frac{1}{2}\varepsilon_0 + \frac{1}{2}\varepsilon_1$ which satisfies (3.1) and not (3.1)'. Conversely, if $V = \{0, 1\} \subset \mathbb{R}$ and μ is Lebesgue measure on [0, 1], (3.1)' holds and (3.1) fails.

One may use theorem 1, under condition (3.1)', to show as in theorem 3 that relatively open convex sets in a Banach space are CS-closed.

This theorem 3 is proven in a number of places (Blackwell and Girshick [1], p. 48, Cook and Webster [4] and ter Morsche [8]) without reference to CS-convexity. What these authors in fact prove is that $conv\{c(n)\}$ is CS-closed. In these terms the result is intrinsically finite dimensional as the following simple proposition shows. Applications of theorem 3 to game theory can be found in Blackwell and Girshick [1], p. 50 and in Tijs [9], pp. 34, 38, 46.

PROPOSITION 2. In any infinite dimensional Banach space X one can find a sequence $\{x_n\}$ with conv $\{x_n\}$ not CS-closed.

Proof. We may pick an infinite dimensional separable closed subspace X_0 of X and a sequence $\{x_n\}$ dense in the unit ball of X_0 . The Baire category theorem shows that $\operatorname{conv}\{x_n\}$ has no non-empty interior in X_0 while $\operatorname{cl conv}\{x_n\}$ certainly has a non-empty interior. CS -closed convex sets in Banach spaces have the same interiors as their closures do (cf. [7], p. 183). Thus $\operatorname{conv}\{x_n\}$ is not CS -closed. \square

We now use theorem 3 to extend a recent result of Cook [2]. Of course, this extension can be seen as a generalization of Carathéodory's theorem.

Let $D = [d_{ij}]$ be an upper bounded or a lower bounded $k \times \infty$ — matrix of real numbers and let $d = (d_1, d_2, \dots, d_k) \in \mathbb{R}^k$. Put $S(D, d) = \{p = (p_1, p_2, \dots) \in S^{\infty} : Dp \in \mathbb{R}^k, Dp \leq d\}$.

Theorem 4. Let x_0, x_1, x_2, \ldots be an infinite sequence in \mathbb{R}^m and let $q = (q_1, q_2, \ldots) \in S(D, d)$ such that $x_0 = \sum_{j=1}^{\infty} q_j x_j$. Then there exists an $r = (r_1, r_2, \ldots) \in S(D, d)$ such that at most m + k + 1 coordinates of r are non-zero and such that $x_0 = \sum_{j=1}^{\infty} r_j x_j$.

Proof. Note that $Dp \in \mathbb{R}^k$ for each $p \in S(D, d)$ if D is lower bounded, and that $Dp \in (\mathbb{R} \cup \{-\infty\})^k$ if D is upper bounded. (a) First suppose that $Dq \in \mathbb{R}^k$.

Then $(x_0, Dq) = \sum_{j=1}^{\infty} q_j(x_j, D_j)$, where D_j is the *j*th column of the matrix D. It follows from theorem 3 (with conv $\{(x_j, D_j)\}$ in the role of C and the (m+k)-dimensional space $\mathbb{R}^m \times \mathbb{R}^k$ in the role of \mathbb{R}^m) that there is an $r \in S^{\infty}$ with at most (m+k)+1 coordinates unequal to zero and such that

$$(x_0, Dq) = \sum_{j=1}^{\infty} r_j(x_j, D_j) = \left(\sum_{j=1}^{\infty} r_j x_j, Dr\right).$$

But then $r \in S(D, d)$ because $Dr = Dq \le d$, and $x_0 = \sum_{j=1}^{\infty} r_j x_j$. Thus we have proved the theorem for the case that $Dq \in \mathbb{R}^k$. (b) Now suppose that $Dq \notin \mathbb{R}^k$. Then D is upper bounded.

Let s be a positive upper bound for the elements of D and let I be the non-empty set $\{i \in \{1, \ldots, k\}: \sum_{j=1}^{\infty} d_{ij}q_j = -\infty\}$. Take a $t \in \mathbb{N}$ such that

$$\sum_{j=1}^{t} d_{ij} q_j \le d_i - s \quad \text{for each} \quad i \in I.$$

Let $E = [e_{ij}]$ be the $k \times \infty$ - matrix with

$$e_{ij} = \max\{0, d_{ij}\}$$
 if $i \in I$ and $j > t$,

and

$$e_{ij} = d_{ij}$$
 otherwise.

Put $S(E, d) = \{p \in S^{\infty} : Ep \in \mathbb{R}^k, Ep \le d\}$. Then it is straightforward to show that $q \in S(E, d)$ and that $Eq \in \mathbb{R}^k$. In view of part (a) of this proof (with E in the role of D) we may conclude that there exists an $r \in S(E, d)$ with at most m + k + 1 coordinates unequal to zero such that $x_0 = \sum_{j=1}^{\infty} r_j x_j$. Now $Dr \le Er \le d$. Hence $r \in S(D, d)$ and we have proved the theorem. \square

Proposition 2 shows that there is no straightforward extension of theorem 4 to infinite dimensional spaces.

- W. D. Cook [2] proved the above theorem under the two additional assumptions:
- (1) The sequence D_1, D_2, \ldots of columns of D is a closed bounded sequence in \mathbb{R}^k .
- (2) x_1, x_2, \ldots is a closed bounded sequence in \mathbb{R}^m .

In his proof he used a duality theorem of semi-infinite programming theory. Our proof is considerably simpler and our result much more general.

Without going into details we note that those results of the paper of Cook, Field and Kirby [3] which were obtained by using Cook's theorem can be strengthened by using theorem 4.

4. An application in mathematical programming theory. Let Y be a set and $m \in \mathbb{N}$. Let f_1, f_2, \ldots, f_m be real-valued lower bounded functions on Y, let f_{m+1} be a real-valued bounded function on Y and let $b = (b_1, b_2, \ldots, b_m) \in \mathbb{R}^m$. By \mathcal{B}

we denote the smallest σ -algebra of subsets of Y such that $f_1, f_2, \ldots, f_{m+1}$ are measurable functions. Let R_1 be the family of those finite measures μ on the measurable space (Y, \mathcal{B}) for which

(4.0)
$$\int f_i(y) d\mu(y) \le b_i \quad \text{for} \quad i = 1, 2, ..., m.$$

Let C be the convex cone generated by the set of probability measures $\{\varepsilon_y : y \in Y\}$, where ε_y is the point measure with mass 1 at y. Let R_2 be the subset of those elements μ of C for which (4.0) holds. For i = 1, 2 we look at

PROBLEM i. Find the value

$$v_i = \inf_{\mu \in R_i} \int f_{m+1}(y) \ d\mu(y)$$

and (if possible) an element of the solution set

$$O_i = \left\{ \mu \in R_i : \int f_{m+1}(y) \ d\mu(y) = v_i \right\}.$$

Note that the problems 1 and 2 coincide if Y is a finite set and that then essentially we have a standard finite linear programming problem.

The following theorem shows that both problems are feasible if one of them is so; that the values of both problems are equal and that both solution sets are non-empty if one of these sets is. Theorem 2 plays a crucial role in the proof of this theorem.

THEOREM 5. With terminology as above let

$$O_i(\delta) = \left\{ \mu \in R_i : \int f_{m+1}(y) \ d\mu(y) \le v_i + \delta \right\}$$

for i = 1, 2 and for each $\delta \ge 0$. Then

$$(4.1) R_1 \neq \emptyset iff R_2 \neq \emptyset$$

$$(4.2)$$
 $v_1 = v_2$

(4.3) for each
$$\delta \geq 0$$
: $O_1(\delta) \neq \emptyset$ iff $O_2(\delta) \neq \emptyset$.

Proof. Since $R_2 \subseteq R_1$, we may conclude that

$$(4.4) R_1 \neq \emptyset if R_2 \neq \emptyset and v_1 \leq v_2.$$

Note that the theorem holds if $R_1 = \emptyset$. Suppose now that we can show that for each $\mu \in R_1$, there is a $\tilde{\mu} \in R_2$ such that

(4.5)
$$\int f_i(y) d\mu(y) = \int f_i(y) d\tilde{\mu}(y) \quad \text{for} \quad i = 1, 2, ..., m + 1.$$

Then we may conclude that (4.1) holds and that $v_2 \le v_1$, and thus $v_2 = v_1$ in

view of (4.4). Furthermore, $\tilde{\mu} \in O_2(\delta)$ if $\mu \in O_1(\delta)$, while it is also obvious that $O_2(\delta) \subset O_1(\delta)$; thus (4.3) holds. Hence, all that remains is the proof of (4.5). Take $\mu \in R_1$. If $\mu(Y) = 0$, then take $\tilde{\mu} = \mu \in R_2$, and (4.5) holds. Suppose now

that $\mu(Y) > 0$. Note that

$$m_i\mu(Y) \leq \int f_i d\mu \leq b_i$$
 for $i = 1, ..., m$

and

$$m_{+1}\mu(Y) \leq \int f_{m+1} d\mu \leq M\mu(Y)$$

where $m_i = \inf_{y \in Y} f_i(y) \in \mathbb{R}$ for $i = 1, \ldots, m+1$ and $M = \sup_{y \in Y} f_{m+1}(y)$. Hence $\int f_i d\mu \in \mathbb{R}$ for $i = 1, 2, \ldots, m+1$. Let $T: Y \to \mathbb{R}^{m+1}$ be the map with $Ty = (f_1(y), f_2(y), \ldots, f_{m+1}(y))$ for each $y \in Y$, and let ρ be the probability measure on \mathbb{R}^{m+1} defined by

$$\rho(A) = (\mu(Y))^{-1}\mu(T^{-1}(A))$$
 for each Borel subset A of \mathbb{R}^{m+1} .

Then, in view of theorem C in Halmos [6] p 163, we have

$$\mu(Y)\int \pi_i(x) d\rho(x) = \int \pi_i(Ty) d\mu(y) = \int f_i(y) d\mu(y) \in \mathbb{R}$$

for i = 1, 2, ..., m + 1. Hence

$$b(\rho) = \int x \, d\rho(x) = (\mu(Y))^{-1} \int Ty \, d\mu(y) \in \mathbb{R}^{m+1}.$$

Let $V = T(T^{-1}(\operatorname{supp}(\rho)))$. Then cl $V = \operatorname{supp}(\rho)$. Hence, in view of theorem 2, there exists a probability measure ν on \mathbb{R}^m with a finite support in V and with $b(\rho) = b(\nu)$. Let $\operatorname{supp}(\nu)$ have k members, then there are $y(1), y(2), \ldots, y(k) \in Y$ such that $\operatorname{supp}(\nu) = \{Ty(1), Ty(2), \ldots, Ty(k)\}$. Moreover, there is a $(p_1, p_2, \ldots, p_k) \in \mathbb{R}^k$, with $p_j \ge 0$ for each $j = 1, 2, \ldots, k$ and $\sum_{j=1}^k p_j = 1$, such that $b(\nu) = \sum_{j=1}^k p_j Ty(j)$. Put $\tilde{\mu} = \mu(Y) \sum_{j=1}^k p_j \varepsilon_{y(j)} \in R_2$. Then

$$\int f_i(y) d\tilde{\mu}(y) = \mu(Y) \sum_{j=1}^k p_j f_i(y(j)) = \mu(Y)(b(\nu))_i = \mu(Y)(b(\rho))_i = \int f_i(y) d\mu(y)$$

for i = 1, 2, ..., m + 1 and thus (4.5) holds. \Box

REMARKS. Theorem 5 may be viewed as an extremal principle for Problem 1. The continuous version of theorem 5 in which Y is a compact Hausdorff space and f_0, f_1, \ldots, f_m are continuous functions in C(Y) shows this more clearly. The theorem then says that equation (4.5) takes the same values on the Borel measures on Y, M(Y), as it does on finite convex combinations of point evaluations in M(Y). These point evaluations are the extreme points of the positive unit ball in $M(Y) = C(Y)^*$. The theorem thus says such extremal combinations suffice.

We want to conclude with an economic interpretation of problems 1 and 2 (cf. Gale [5], pp. 5–8). We look at a linear production system in which m goods G_1, G_2, \ldots, G_m are involved and where Y is the set of possible activities. We suppose that there is a fixed supply b_i of goods G_i (i = 1, 2, ..., m) and that activity $y \in Y$ operated at intensity 1 needs $f_1(y), f_2(y), \ldots, f_m(y)$ units of goods G_1, G_2, \ldots, G_m , respectively. Then elements of R_1 can be seen in this context as feasible production schedules and elements of R₂ as finite feasible schedules e.g. production schedules using only a finite number of activities. Let $-f_{m+1}(y)$ be the income associated with activity y at intensity 1. Then problem 1 (problem 2) corresponds with the problem of finding a production schedule (a finite production schedule) which maximizes the total income without exceeding the given supplies. Now it follows from (4.5) that for each feasible $\mu \in R_1$ there is a feasible production schedule $\tilde{\mu}$ as good as μ , while $\tilde{\mu}$ uses only a finite number of activities. Thus we may conclude that we can restrict our attention to production schedules in which only a finite number of activities are involved. Note that even m+1 activities suffice.

It will be obvious that in other economic situations theorem 5 may also be useful.

ACKNOWLEDGEMENT. We thank the referee for valuable remarks which improved the presentation of the material in this paper considerably.

REFERENCES

- 1. D. Blackwell and M. A. Girshick, Theory of games and statistical decisions, John Wiley, New York, 1954.
- 2. W. D. Cook, Carathéodory's Theorem with Linear Constraints, Canadian Math. Bull. 17 (1974), pp. 189-191.
- 3. W. D. Cook, C. A. Field and M. J. L. Kirby, Infinite Linear Programming in Games with Partial Information, Operations Research 23 (1975), pp. 996–1010.
- 4. W. D. Cook and R. J. Webster, Carathéodory's Theorem, Canadian Math. Bull. 15 (1972), p. 293.
 - 5. D. Gale, The Theory of Linear Economic Models, McGraw-Hill, New York, 1960.
 - 6. P. R. Halmos, Measure Theory, van Nostrand, Princeton (6th printing) 1959.
 - 7. G. Jameson, Ordered Linear Spaces, Springer, Berlin, 1970.
- 8. H. G. ter Morsche, *Problem* 341, Problem Section Nieuw Archief voor Wiskunde **21** (1973), p. 98. See also: *Solutions of problem* 341, Nieuw Archief voor Wiskunde **21** (1973), pp. 292–293 and **22** (1974), p. 85.
- 9. S. H. Tijs, Semi-infinite and infinite matrix games and bimatrix games, Ph.D. dissertation, University of Nijmegen, The Netherlands, 1975.

DEPARTMENT OF MATHEMATICS
CATHOLIC UNIVERSITY
NIJMEGEN, THE NETHERLANDS

DEPARTMENT OF MATHEMATICS
DALHOUSIE UNIVERSITY
HALIFAX, NOVA SCOTIA
CANADA
(Partially supported on
N.R.C. A4493)